EFFICIENT TRAINING OF 3D UNROLLED NEURAL NETWORKS FOR MRI RECONSTRUCTION USING SMALL DATABASES

Zilin Deng[1,2], Burhaneddin Yaman[1,2], Chi Zhang[1,2], Steen Moeller[2], and Mehmet Akçakaya[1,2]

[1]Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN
[2]Center for Magnetic Resonance Research, University of Minnesota, Minneapolis, MN

ABSTRACT

3D MRI encodes volumetric information, typically offering improved contiguous coverage and resolution than 2D MRI. However, 3D MRI data acquisition is lengthy, and requires accelerated imaging techniques. Deep learning methods have recently emerged as a powerful strategy for MRI reconstruction. Among such methods, unrolled networks have proven powerful with their ability to incorporate the forward encoding operator directly. These methods are largely applied in a 2D setting, but 3D processing has the potential to further improve reconstruction quality for volumetric imaging by capturing multi-dimensional interactions. Nevertheless, implementing 3D unrolled networks is challenging because of memory limitations on GPUs, as well as the lack of large databases of 3D k-space data. To tackle both of these issues, we propose a data augmentation strategy that generates smaller sub-volumes from large volumetric datasets. Subsequently, these augmented datasets are used to train a 3D unrolled network, and compared to their 2D counterpart. The results show that our 3D processing provides improved reconstruction results on volumetric data than 2D processing.

***Index Terms*—** 3D processing, deep learning, network training, algorithm unrolling, MRI

1. INTRODUCTION

In several clinical applications, 3D magnetic resonance imaging (MRI) provides extensive contiguous coverage, higher signal-to-noise ratio (SNR) and improved spatial resolution to generate improved depiction of small structures by multiplanar reformatting [1, 2]. Nevertheless, long acquisition times hinder the utility of 3D MRI in clinical workflows [3]. One solution is to reduce the scan time of such acquisitions by applying accelerated imaging techniques, where sub-sampled data are reconstructed with additional information. Numerous accelerated MRI methods have been proposed, including parallel imaging [4, 5] that utilizes redundant information from receiver coils, and compressed sensing that uses the compressibility of MR images.

Deep learning (DL) methods have recently emerged as an efficient strategy for accelerated MRI reconstruction [6–14]. Among the proposed DL methods, physics-guided (PG) approaches are a category that successfully produce high-quality reconstructions of undersampled MRI data, while incorporating the physics of the MRI acquisition via the explicit use of the forward encoding operator. In unrolled networks [12], conventional iterative algorithms for solving a regularized least squares objective function, which alternate between enforcing data consistency (DC) and performing a proximal operation based on the regularizer, are unrolled for a fixed number of iterations. The DC unit is solved using standard linear methods, while the regularization units are implemented via neural networks most of which are based on 2D convolutions [6, 7, 10–12].

When applying unrolled networks to 3D MRI reconstruction, a common strategy is to divide the volumetric MRI data to 2D slices, which are then processed by unrolled networks with 2D convolutions. Nevertheless, for volumetric 3D imaging, 3D processing may further improve reconstruction results compared to 2D processing due to its ability to capture multi-dimensional interactions [15]. Some approaches using 3D convolutional kernels have been proposed, which either apply specialized training tools [9] or simplify the data consistency approach [13]. However, it remains challenging to train unrolled networks with 3D processing. Large volumetric MRI data is hard to handle in GPUs due to the memory limitations. On the other hand, the scarcity of 3D raw k-space datasets may cause overfitting problem during the training of the network.

In this study, we tackle these two issues by generating multiple 3D slabs of smaller size from the full 3D Cartesian MRI volume as a means of data augmentation. Results on 3D knee MRI data [16] show that our 3D processing provides improved performance compared to 2D processing on data with high acceleration rates, due to its ability to capture multi-dimensional information.

2. MATERIALS AND METHODS

2.1. Unrolled Networks for MRI reconstruction

The inverse problem for accelerated MRI reconstruction can be formulated as a regularized least-squares problem:

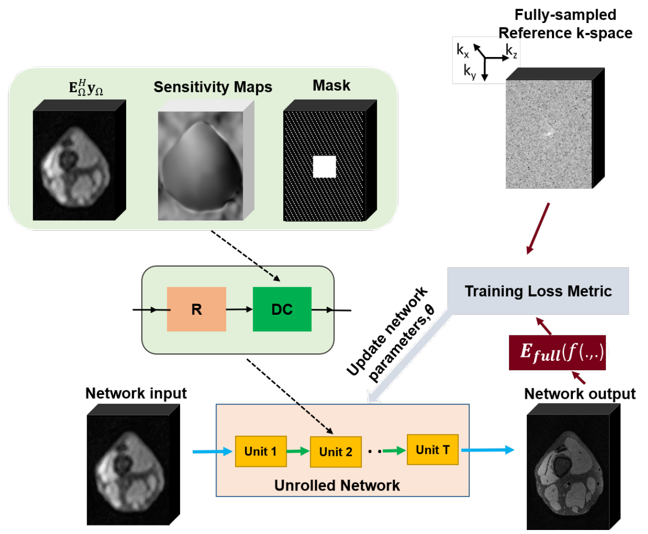

Fig. 1: Schematic of the supervised training paradigm for an unrolled MRI reconstruction network. In the unrolled network, each unrolled unit contains a regularizer (R), which is implicitly solved via a neural network, and a data consistency (DC) component solved by linear methods.

$$\arg\min_{\mathbf{x}} \|\mathbf{y}_\Omega - \mathbf{E}_\Omega \mathbf{x}\|_2^2 + \mathcal{R}(\mathbf{x}), \quad (1)$$

where $\mathbf{x}$ is the image of interest, $\mathbf{y}_\Omega$ is the acquired measurements with sub-sampling pattern Ω, $\mathbf{E}_\Omega : \mathbb{C}^M \to \mathbb{C}^P$ is the multi-coil encoding operator containing coil sensitivities and partial Fourier matrix sampling, and $\mathcal{R}(\cdot)$ is a regularizer. Thus, the first quadratic term in the objective function enforces DC, while the second term is for regularization. Conventionally, when using an explicit regularizer, this is typically solved via an iterative algorithm [17].

Variable splitting with quadratic penalty (VSQP) [17] is an iterative algorithm commonly used to solve such inverse problem. In VSQP, an auxiliary variable $\mathbf{z}$ is introduced to decouple the DC and the regularizer. The inverse problem in (1) is then transformed to

$$\arg\min_{\mathbf{x}} \|\mathbf{y}_\Omega - \mathbf{E}_\Omega \mathbf{x}\|_2^2 + \mu\|\mathbf{x} - \mathbf{z}\|_2^2 + \mathcal{R}(\mathbf{x}), \quad (2)$$

where μ is the penalty parameter. This is solved by alternating minimization over $\mathbf{z}$ and $\mathbf{x}$ as

$$\mathbf{z}^{(i)} = \arg\min_{\mathbf{z}} \mu\|\mathbf{x}^{(i-1)} - \mathbf{z}\|_2^2 + \mathcal{R}(\mathbf{z}) \quad (3a)$$

$$\mathbf{x}^{(i)} = \arg\min_{\mathbf{x}} \|\mathbf{y}_\Omega - \mathbf{E}_\Omega \mathbf{x}\|_2^2 + \mu\|\mathbf{x} - \mathbf{z}^{(i)}\|_2^2 \quad (3b)$$

where $\mathbf{z}^{(i)}$ is an intermediate output and $\mathbf{x}^{(i)}$ is the desired image at iteration i.

In algorithm unrolling [12], such conventional iterative algorithms are unrolled for a fixed number of iterations. Each iteration contains a regularizer unit, which is solved with neural networks, and a DC which can be solved with standard techniques such as the conjugate gradient (CG) method [7]. Figure 1 shows the training process, where the unrolled network is trained end-to-end in a supervised manner by minimizing the loss defined as

$$\min_{\boldsymbol{\theta}} \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}(\mathbf{y}_{\text{ref}}^i,\ \mathbf{E}_{\text{full}}^i f(\mathbf{y}_\Omega^i, \mathbf{E}_\Omega^i; \boldsymbol{\theta})), \quad (4)$$

where N is the number of samples in the training database, $\mathcal{L}(\cdot,\cdot)$ is a loss function, $\mathbf{y}_{\text{ref}}^i$ is the fully-sampled k-space for subject i, $f(\mathbf{y}_\Omega^i, \mathbf{E}_\Omega^i; \boldsymbol{\theta})$ is the output of the unrolled network for sub-sampled k-space data $\mathbf{y}_\Omega^i$ with the network being parameterized by $\boldsymbol{\theta}$, and $\mathbf{E}_{\text{full}}^i$ is the fully-sampled encoding operator that transforms network output to k-space.

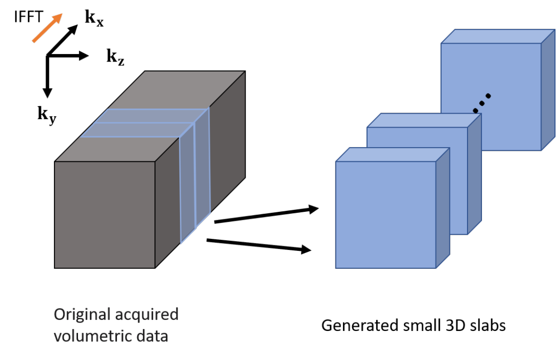

Fig. 2: The proposed data augmentation strategy for generating small sub-volumes from volumetric datasets. IFFT is applied in the fully sampled (k_x) direction, and the resulting hybrid image-frequency volume is divided into multiple sub-volumes.

2.2. 3D MRI Data and Database Augmentation

Unrolled networks are mostly implemented with 2D convolutions [6, 7, 10–12]. A common strategy to apply unrolled networks in 3D Cartesian MRI reconstruction is to divide the volumetric data into 2D slices by taking the inverse Fourier transform along the fully-sampled readout direction, which are then reconstructed with 2D processing. Recently, 3D kernels have also been used in some studies, either through specialized training tools [9] or simplified DC approaches [13]. However, it remains challenging to train unrolled networks with 3D processing due to both memory constraints of GPUs and the lack of availability of large databases of 3D data, even though 3D processing has the potential to offer improved reconstruction quality compared to its 2D counterpart.

We tackle these two issues by applying a data augmentation strategy to generate multiple 3D slabs of smaller size from the full 3D volume. An overview of this data augmentation strategy for the generation of the sub-volumes is summarized in Figure 2. First, the inverse Fourier transform of the acquired undersampled k-space data is taken along the fully-sampled frequency encoding (k_x) direction. Then, the hybrid image-frequency volume is divided into smaller, potentially

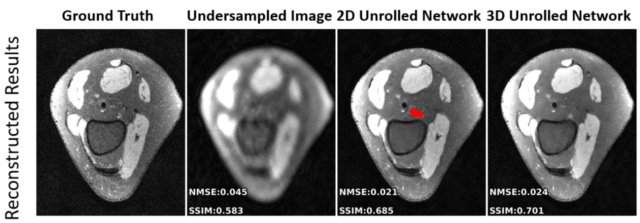

Fig. 3: A representative test slice from reconstructions using 2D and 3D unrolled networks for an acceleration rate of 8. The red arrow shows a residual aliasing artifact present in the 2D processed reconstruction, which is removed with 3D processing.

overlapping, 3D slabs in this direction, which are then transformed back to k-space. This approach enables the generation of 310 smaller 3D slabs from as few as 10 subjects out of a database of 20 fully-sampled 3D knee MRI datasets [16]. In addition to augmenting the training database to have more training samples, the use of the smaller slabs also leads to a lower GPU memory footprint, which facilitates the training of unrolled networks with standard methods.

2.3. Experiments and Evaluation

Our proposed data augmentation and training strategy was tested by retrospectively undersampling the fully-sampled k-space in the k_y-k_z plane for 3D knee MRI data from the database in [16]. The database consisted of 20 subjects, who were scanned at 3T (8-channel coil array) with FOV $= 160 \times 160 \times 154\ \mathrm{mm}^3$, resolution $= 0.5 \times 0.5 \times 0.6\ \mathrm{mm}^3$, matrix size $= 320 \times 320 \times 256$. The undersampling was performed at acceleration rates 8 and 12 with ACS $= 32 \times 32$ using a sheared uniform undersampling pattern in k_y-k_z [18].

The 3D unrolled network consisted of 5 unrolled blocks, each including a 3D ResNet [14] architecture comprising 5 residual blocks with $3 \times 3 \times 3$ convolutional kernels for the regularizer unit. The data consistency problem was solved using a conjugate gradient approach that was unrolled for 5 iterations [7]. The 3D approach was compared to a 2D unrolled network, which was designed in the same way, except for using 2D convolutions with 3×3 kernels and 15 residual blocks in the ResNet instead of 3D convolutions, which led to the same number of trainable parameters. The 3D network was trained using 310 small slabs while the 2D network was trained using 310 slices generated from the volumetric data. Both networks are trained end-to-end minimizing the normalized ℓ_1-ℓ_2 loss [11, 19] defined as

$$\mathcal{L}(\mathbf{u}, \mathbf{v}) = \frac{\|\mathbf{u} - \mathbf{v}\|_2}{\|\mathbf{u}\|_2} + \frac{\|\mathbf{u} - \mathbf{v}\|_1}{\|\mathbf{u}\|_1}. \quad (5)$$

Both the 3D and 2D unrolled networks were trained for 100 epochs with a learning rate of 5×10^{-4}. Testing for 3D processing was performed on 10 different subjects from [16]. The reconstructions were quantitatively evaluated using normalized mean square error (NMSE) and structural similarity index (SSIM).

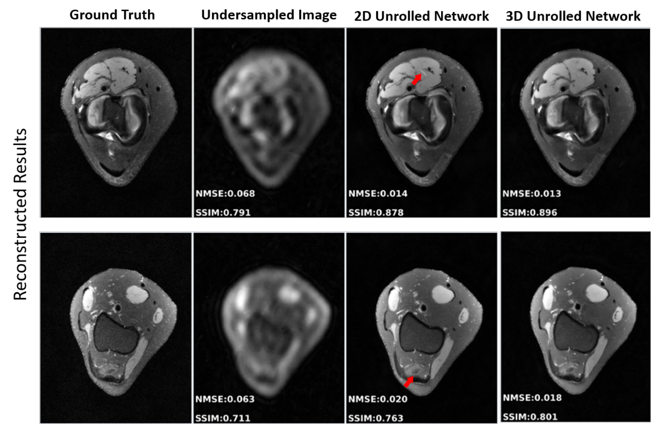

Fig. 4: Two representative test slices from reconstructions using 2D and 3D unrolled networks for an acceleration rate of 12. Red arrows indicate cases where 3D processing reduces residual artifacts (top row) and preserves fine details (bottom row).

3. RESULTS

Figure 3 shows reconstruction results on a representative test slice with acceleration rate 8. 2D processing suffers from residual artifacts caused by undersampling (red arrow), while 3D processing removes such artifacts. Same observations apply to Figure 4, which shows testing results for acceleration rate 12. Artifacts shown by the red arrows in 2D processing are successfully removed in 3D processing. Furthermore, 3D processing provides improved details compared to 2D processing.

Table 1 displays the quantitative evaluation of the reconstruction methods, where the proposed 3D processing leads to higher SSIM and lower NMSE, indicating that the way 3D processing captures multi-dimensional interactions compared

Table 1: Median and interquartile ranges [25^{th}-75^{th} percentile] of SSIM and NMSE metrics on the 3D knee MRI test dataset at acceleration rates (R) of 8 and 12.

Quantitative Metric	2D Unrolled Network	3D Unrolled Network
SSIM (R=8)	0.8280 [0.7656, 0.8690]	0.8381 [0.7782, 0.8725]
NMSE (R=8)	0.0140 [0.0120, 0.0178]	0.0138 [0.0119, 0.0170]
SSIM (R=12)	0.8072 [0.7401, 0.8503]	0.8277 [0.7653, 0.8633]
NMSE (R=12)	0.0169 [0.0149, 0.0208]	0.0164 [0.0144, 0.0196]

to 2D processing further benefit the reconstruction quality.

4. DISCUSSION AND CONCLUSION

Unrolled deep neural networks have shown great promise in improving accelerated MRI reconstruction. However, most of these networks are applied with 2D processing. To apply unrolled networks for 3D MRI reconstruction, a common strategy is to divide the volumetric MRI data to 2D slices, which are then processed with 2D convolutions, which nevertheless loses multi-dimensional interactions. On the other hand, directly training 3D unrolled networks with 3D raw k-space data faces problems due to the lack of availability of large databases of 3D datasets and GPU memory limitations. In this work, we proposed a data augmentation strategy for 3D processing that enables efficient training of 3D unrolled networks for volumetric MRI reconstruction. Our proposed data augmentation strategy generated smaller 3D subvolumes from large Cartesian volumetric acquisitions, which helped tackle both these challenges without needing advanced training strategies [9]. Our results show that the 3D unrolled networks have the potential to improve MRI reconstruction at high acceleration rates by capturing multi-dimensional interactions.

References

[1] R. Kijowski and G. E. Gold, "Routine 3D magnetic resonance imaging of joints," *Journal of Magnetic Resonance Imaging*, vol. 33, no. 4, pp. 758–771, 2011.

[2] R. M. Botnar, M. Stuber, P. G. Danias, K. V. Kissinger, and W. J. Manning, "Improved coronary artery definition with T2-weighted, free-breathing, three-dimensional coronary MRA," *Circulation*, vol. 24, pp. 3139–3148, 1999.

[3] A. Naraghi and L. M. White, "Three-Dimensional MRI of the Musculoskeletal System," *American Journal of Roentgenology*, vol. 199, no. 3, pp. W283–W293, 2012.

[4] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Magn Reson Med*, vol. 42, pp. 952–962, 1999.

[5] M. A. Griswold, P. M. Jakob, et al., "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn Reson Med*, vol. 47, pp. 1202–1210, 2002.

[6] K. Hammernik, T. Klatzer, et al., "Learning a variational network for reconstruction of accelerated MRI data," *Magn Reson Med*, vol. 79, pp. 3055–3071, 2018.

[7] H. K. Aggarwal, M. P. Mani, and M. Jacob, "MoDL: Model-Based Deep Learning Architecture for Inverse Problems," *IEEE Trans Med Imaging*, vol. 38, pp. 394–405, 2019.

[8] F. Knoll, T. Murrell, et al., "Advancing machine learning for MR image reconstruction with an open competition: Overview of the 2019 fastMRI challenge," *Magn Reson Med*, Jun 2020.

[9] M. Kellman, K. Zhang, et al., "Memory-efficient learning for large-scale computational imaging," *IEEE Trans Comp Imaging*, vol. 6, pp. 1403–1414, 2020.

[10] S. A. H. Hosseini, B. Yaman, S. Moeller, M. Hong, and M. Akçakaya, "Dense recurrent neural networks for accelerated mri: History-cognizant unrolling of optimization algorithms," *IEEE J Sel Top Sig Proc*, vol. 14, no. 6, pp. 1280–1291, 2020.

[11] B. Yaman, S. A. H. Hosseini, et al., "Self-Supervised Learning of Physics-Guided Reconstruction Neural Networks without Fully-Sampled Reference Data," *Magn Reson Med*, vol. 84, no. 6, pp. 3172–3191, Dec 2020.

[12] F. Knoll, K. Hammernik, et al., "Deep-learning methods for parallel magnetic resonance imaging reconstruction: A survey of the current approaches, trends, and issues," *IEEE Sig Proc Mag*, vol. 37, no. 1, pp. 128–140, 2020.

[13] T. Küstner, N. Fuin, et al., "Cinenet: deep learning-based 3D cardiac cine MRI reconstruction with multi-coil complex-valued 4D spatio-temporal convolutions," *Scientific Reports*, 2020.

[14] B. Yaman, C. Shenoy, Z. Deng, S. Moeller, and M. Akçakaya, "Self-supervised physics-guided deep learning reconstruction for high-resolution 3d lge cmr," in *Proc. IEEE Int. Symp. Biomed. Imag. (ISBI)*. IEEE, 2021.

[15] B. Yaman, S. Weingartner, N. Kargas, N. D. Sidiropoulos, and M. Akcakaya, "Low-Rank Tensor Models for Improved Multidimensional MRI: Application to Dynamic Cardiac T_1 Mapping," *IEEE transactions on computational imaging*, vol. 6, pp. 194–207, 2019.

[16] F. Ong, S. Amin, S. Vasanawala, and M. Lustig, "Mridata. org: An open archive for sharing mri raw data," in *Proc. Intl. Soc. Mag. Reson. Med*, 2018, vol. 26, p. 1.

[17] J. A. Fessler, "Optimization methods for magnetic resonance image reconstruction: Key models and optimization algorithms," *IEEE Sig Proc Mag*, vol. 37, no. 1, pp. 33–40, 2020.

[18] F. A. Breuer, M. Blaimer, et al., "Controlled aliasing in volumetric parallel imaging (2D CAIPIRINHA)," *Magn Reson Med*, vol. 55, no. 3, pp. 549–556, Mar 2006.

[19] B. Yaman, S. Hosseini, et al., "Multi-mask self-supervised learning for physics-guided neural networks in highly accelerated MRI," *arXiv preprint arXiv:2008.06029*, 2020.